

Effective Atomic Charges in Alanine Dipeptide

J. RIDARD, B. LÉVY

Laboratoire de Physico-Chimie des Rayonnements (UA CNRS 75), Bât 350, Université Paris Sud, 91405 Orsay Cedex, France

Received 28 July 1998; accepted 23 October 1998

ABSTRACT: A new set of effective atomic charges of different conformers of alanine dipeptide is presented. These charges are obtained by fitting the electrostatic potential resulting from the *ab initio* SCF wave function of the system obtained in a 6-31G basis set. A specific fit procedure is used providing charges weakly dependent on the fit points as well as on the geometry of the molecule. It is shown that these charges retain a reasonable chemical meaning. © 1999 John Wiley & Sons, Inc. J Comput Chem 20: 473–482, 1999

Keywords: atomic charges; *ab initio* SCF calculation; electrostatic potential; least square fit; alanine dipeptide

Introduction

In this study, we re-examine the determination of effective atomic charges of alanine dipeptide (Fig. 1). An efficient procedure to determine such charges consists of fitting the electrostatic potential created by the considered molecule and evaluated by an *ab initio* method.

This procedure has been studied by many investigators and recent reviews are available.^{1–3} The main difficulty comes from the fact that charges determined in this way often depend on the fit points and on the geometry of the molecules,⁴ particularly for alanine dipeptide.^{5,6}

Several techniques have been proposed to overcome this difficulty: better point selection^{3,6}; re-

duction in rank deficiency of the problem by selecting a subset of charges to be evaluated^{7,8}; and a determination of charges by fitting the moments of the electronic charge distribution instead of fitting them on the electrostatic potential.^{3,9}

Here, we present the results obtained for alanine dipeptide by using a new method,¹⁰ which has proven to be efficient in providing stable effective atomic charges; that is, charges that are weakly dependent on the fit points¹⁰ as well as on the geometry, particularly in the case of aliphatic chains.¹¹

Method

The technique used here to determine stable effective atomic charges is summarized in what follows.

Correspondence to: B. Lévy

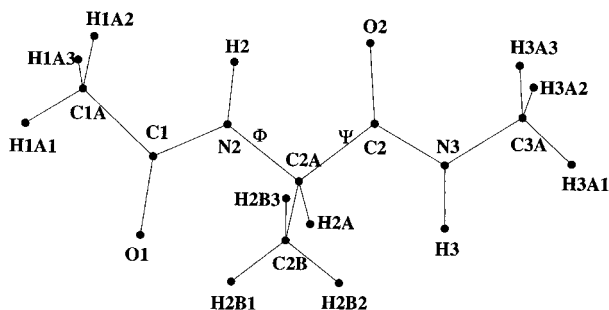


FIGURE 1. Projection of alanine dipeptide (conformer II) in the N2—C2A—C2 plane showing the labels of the atom. Note that the H2A atom has been moved slightly to the right and the methyl group with the C2B carbon atom to the left for clarity. In actual projection, the four atoms (C2A, H2B3, H2A and C2B) would lie on the same straight line.

Charges are determined by a least-squares fit of an *ab initio* electrostatic potential. This leads to a set of linear equations that can be written in matrix form as:

$$\mathbf{A}\mathbf{q} = \mathbf{B} \quad (1)$$

where \mathbf{A} is a square matrix, and \mathbf{B} and \mathbf{q} are column vectors. The element q_k of \mathbf{q} is the unknown effective charge of the atom k and the elements of \mathbf{A} and \mathbf{B} are defined as:

$$A_{kl} = \sum_i \frac{1}{r_{ik} r_{il}} \quad (2)$$

$$B_k = \sum_i V_i \frac{1}{r_{ik}}$$

where the following notation is used: k, l the indices of the atoms of the system; i is the index of a fit point; r_{ik} is the distance between the point i and the atom k ; and V_i is the value of the *ab initio* potential at point i .

The solution of eq. (1) is, in principle, straightforward. However, some care has to be exercised because \mathbf{A} is often singular or nearly singular.^{4,8,10} This singularity is easily understood if all the fit points stand at large distances from the molecule. Indeed, in such a case, only the total charge (or dipole) of the molecule contributes to the electrostatic potential and only one (or two) effective charges can be determined.^{10,11} Numerically, this means that all the eigenvalues of \mathbf{A} except one (or two) vanish.

If all the fit points stand at shorter distances the problem should be better conditioned and it should

be possible to determine more than one or two charges. But, experience shows that one must go down to very short distances to get an \mathbf{A} matrix that is not singular, with such distances being too small to have any meaning.

Therefore, the \mathbf{A} matrix does have several small eigenvalues in any useful case, and these small eigenvalues result in numerical instabilities of the charges: The charges determined by solving eq. (1) depend significantly of irrelevant details of the calculation (choice of fit points, geometry changes, etc.).

We suggest the following procedure to overcome this instability.¹⁰ One considers a subspace \mathcal{E} , of the vector space where \mathbf{A} is acting: \mathcal{E} is defined as being spanned by the eigenvectors of \mathbf{A} corresponding to the eigenvalues that are larger than a given threshold, ϵ . Eq. (1) is then solved in \mathcal{E} , not in the whole space, and one then adds the projection of the Mulliken charges into the subspace complementary of \mathcal{E} . It leads to:

$$\mathbf{q} = (\mathbf{PAP} + \mathbf{Q})^{-1} \mathbf{PB} + \mathbf{Qq}_{\text{Mullik}} \quad (3)$$

where $\mathbf{q}_{\text{Mullik}}$ is the column vector of the Mulliken charges and \mathbf{P} and \mathbf{Q} are projection operators; \mathbf{P} projects into \mathcal{E} and \mathbf{Q} into the complementary subspace.

In this way, a charge is determined for all the atoms and a stabilization is obtained by the facts that the $\mathbf{PAP} + \mathbf{Q}$ matrix is nonsingular by definition, that the Mulliken charges are independent of the fit points, and that they have no spurious variations with geometry. In addition, the Mulliken charges correspond usually to chemical intuition (see a more specific discussion of this last point later in this study).

Of course the Mulliken charges do not provide, by themselves, a good fit of the potential. But here, they have a limited influence on the fit (no influence if threshold ϵ vanishes) due to the \mathbf{Q} operator acting on $\mathbf{q}_{\text{Mullik}}$ in eq. (3); that is, they are used merely to give some meaning to the final effective charges.

In addition, it can be shown that \mathbf{q} satisfies eq. (1) exactly if $\epsilon = 0$, whatever the dimension of \mathcal{E} . Otherwise, \mathbf{q} satisfies eq. (1) approximately, and the error increases with ϵ .

Therefore, there are two requirements for this stabilization technique to work properly:

1. The fit points should stand at a distance from the molecule large enough to ensure that the dimension of \mathcal{E} can be large (for purpose of

good stabilization) with a small value of ϵ (for purpose of small error, by neglecting the contribution coming from the subspace complementary of the \mathcal{E} subspace).

2. The fit points should still not be too far from the molecule in order to cover the chemically meaningful region.

Note that eqs. (1) and (3) have been modified slightly (by removing one line and one column in the **A** matrix) to take into account the requirement that the total charge of the system is zero.

We have used both eq. (1) and eq. (3) to evaluate effective atomic charges of alanine dipeptide in a series of 12 different conformers. These conformers are the same as the ones considered in ref. 5. They differ from each other only by the values of the torsion angles around the C2A—N2 and C2A—C2 bonds (these angles are noted Φ and Ψ , respectively (see Fig. 1); the values of Φ and Ψ for the 12 considered conformers are given in Table I). All the remaining internal coordinates are frozen to standard values (see ref. 5 for details).

For each conformer we have performed an *ab initio* SCF calculation using a 6-31G basis set. The GAMESS program system¹² was used to perform this calculation and then to calculate the electrostatic potential created by the nuclei and the electronic cloud. The SCF energy of each conformer relative to conformer I are given in Table I; these energies coincide essentially with the ones of ref. 5.

TABLE I.
Torsion Angles Φ and Ψ (See Fig. 1) and SCF Energies ΔE (Relative to conformer I) of 12 Conformers of Alanine Dipeptide Considered.

Conformer	Φ	Ψ	ΔE (kcal/mol)
I	−84.6	73.0	0.0
II	−165.7	167.3	0.36
III	74.6	−62.0	10.98
IV	−134.2	38.1	4.20
V	−91.9	−5.5	5.33
VI	60.8	40.6	9.54
VII	−161.7	−55.4	7.84
VIII	−112.9	165.4	3.45
IX	−57.0	−47.0	8.22
X	57.0	47.0	9.66
XI	−119.0	113.0	1.94
XII	−139.0	135.0	1.13

The grid of points where the potential was evaluated was obtained by using the keyword PTSEL = CONNOLLY in the GAMESS system. This grid consists of points standing on a series of fused spheres centered on the atoms.

Note that, in this case, no correlation effect was taken into account in spite of the fact that it is probably nonnegligible¹³: the main interest here is the effect of replacing eq. (1) with eq. (3) (i.e., the stabilization of the charges). We assume that this effect is independent of electronic correlation.

Choice of Fit Points

We first study how the effective charges obtained using eq. (1) depend on the fitting grid and how this dependence is reduced by using eq. (3). To that end, we have used different values of r_{\min} , the radius of the innermost spheres used in GAMESS to define the fit points. Let us write:

$$r_{\min} = \text{scal } r_{\text{vdW}} \quad (4)$$

where r_{vdW} is the van der Waals radius of the atom. Here, we have used $\text{scal} = 1.4$ (default value in GAMESS) and $\text{scal} = 2$ [a value for which the charges resulting from eq. (1) are close to those in ref. 5]. The number of fit points obtained in this way varies with scal and with the conformer; that is, it ranges from 959 to 1044 for $\text{scal} = 1.4$ and from 1350 to 1415 for $\text{scal} = 2$.

It is first seen in Table II (columns *a* and *b*) that the charges obtained for conformer I with $\text{scal} = 2$ and eq. (1) are reasonably close to those given in ref. 5. The largest difference is equal to 0.027 electron (atom C2B).

It is also seen in Table II (columns *b* and *c*) that the charges obtained with $\text{scal} = 1.4$ and $\text{scal} = 2$ are significantly different: by 0.07–0.1 electron for most atoms and up to 0.23 electron for C3A. Such differences clearly point out that there is a problem in choosing the fit points when using eq. (1), and that some stabilization process is desirable.

We turn now to the charges obtained using eq. (3). In this case, one has to choose the value of the threshold, ϵ . Here we use:

$$\epsilon = 10^{-4} a_{\max} \quad (5)$$

where a_{\max} is the largest eigenvalue of the **A** matrix. We have found empirically that this choice

TABLE II.
Effective Atomic Charges in Alanine Dipeptide (Conformer I).

Atoms	Computational method					
	Ref. 5	Eq. (1) <i>scal</i> = 2.0	Eq. (1) <i>scal</i> = 1.4	Eq. (3) <i>scal</i> = 2.0	Eq. (3) <i>scal</i> = 1.4	Mulliken
	<i>a</i>	<i>b</i>	<i>c</i>	<i>d</i>	<i>e</i>	<i>f</i>
C2A	.296	.298	.283	−.001	−.003	−.064
H2A	.029	.031	.042	.123	.126	.234
N2	−.683	−.680	−.755	−.668	−.677	−.844
H2	.351	.351	.378	.367	.371	.381
C1	.860	.863	.912	.958	.937	.749
O1	−.729	−.731	−.732	−.761	−.750	−.644
C2	.711	.695	.714	.930	.936	.801
O2	−.670	−.665	−.665	−.697	−.700	−.633
N3	−.504	−.494	−.563	−.730	−.740	−.885
H3	.364	.367	.367	.425	.426	.419
C1A	−.396	−.422	−.491	−.517	−.518	−.533
H1A1	.135	.142	.158	.165	.168	.219
H1A2	.105	.108	.129	.128	.132	.187
H1A3	.102	.111	.128	.128	.129	.189
C3A	−.498	−.503	−.270	−.193	−.196	−.199
H3A1	.199	.196	.142	.129	.132	.163
H3A2	.197	.199	.140	.120	.127	.181
H3A3	.196	.198	.136	.121	.121	.184
C2B	−.403	−.430	−.392	−.444	−.440	−.431
H2B1	.122	.130	.116	.142	.143	.188
H2B2	.098	.104	.102	.132	.137	.155
H2B3	.119	.127	.120	.141	.139	.183
RMS	0.29	0.29	1.04	0.33	1.10	2.74
(kcal/mol)						
RRMS (%)	3.8	3.8	8.0	4.4	8.5	36.0

of ϵ combined with the value $scal = 2$ leads to an **A** matrix with a sufficient number of small eigenvalues and thus to good stabilization.

Thus, it is seen in Table II (columns *d* and *e*) that the charges are now nearly independent of $scal$; that is, the difference between the charges obtained with the two values of $scal$ is now less than 0.02 electron for all atoms. This efficient stabilization is obtained at the expense of accuracy, but to a very minor extent: the root-mean-square (RMS) deviation between the *ab initio* potential and the potential resulting from the effective atomic charges increases by less than 0.1 kcal/mol for each value of $scal$ when passing from eq. (1) to eq. (3). Furthermore, the relative RMS (RRMS) deviation increases by less than 1%.

We conclude that the problem due to the dependence of the charges on the choice of fit points is actually solved by using eq. (3).

Geometry Changes

We now study the way in which the charges vary in the series of the 12 conformers defined in Table I. Only the charges obtained with $scal = 2$ are considered.

The stabilization of these variations obtained by using eq. (3) instead of eq. (1) is best illustrated by Figure 2. This figure concerns the central carbon atom C2A and the four adjacent atoms [the atoms exhibiting the largest geometry dependence of the charges calculated with eq. (1)]. It shows very clearly that the amplitude of the variations of the charges among the conformers is much smaller when using eq. (3) than when using eq. (1), the latter being very close to the amplitude found in ref. 5. A systematic shift of the charges also appears (except in the case of N2).

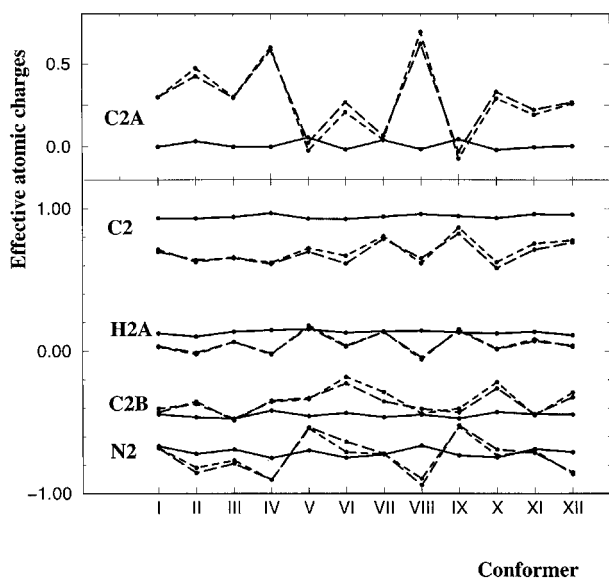


FIGURE 2. Effective atomic charges of C2A and of the four adjacent atoms in the 12 conformers examined. The charges are obtained by the following methods: solid lines: eq. (3) (stabilized charges); long-dashed lines: eq. (1) (nonstabilized charges), and dashed lines: ref. 5.

The degree of stabilization with respect to geometry obtained using eq. (3) can be also characterized by considering the largest and smallest effective charge, q_{\max} and q_{\min} , respectively, of each atom in the different conformers. These values are given in Table III.

Let us first consider the two terminal methyl groups and the atoms O1, O2, N3, and H3. In these atom groups, the differences $q_{\max} - q_{\min}$ are small for all atoms and for the three methods [Mulliken, eq. (3) and eq. (1)]. Thus, no stabilization of the results obtained in eq. (1) appears necessary here. The only exception is the C1A atom, for which the difference $q_{\max} - q_{\min}$, evaluated by eq. (1), is surprisingly large.

The situation is different for the ten remaining atoms. The values of $q_{\max} - q_{\min}$ corresponding to the Mulliken charges and to eq. (3) are small (smaller than 0.1 electron, except for H2B3, for which it rises to 0.113) and close to each other (difference smaller than 0.06 electron), whereas the values obtained by solving eq. (1) are about four times larger than the value obtained by eq. (3) and even more for C2A, N2, C2, and C2B.

The values of $q_{\max} - q_{\min}$ obtained here by using eq. (3) are much smaller than the values in refs. 5 and 6. For instance, the value found for $q_{\max} - q_{\min}$ in ref. 5 for the central C2A atom is equal to 0.764 electron; in Ref. 6, a sophisticated

choice of fit points reduces this value to 0.47–0.54 electron (depending of the points selection); whereas the value obtained here, using the eq. (3), is equal to 0.074 electron.

Eq. (3) thus appears to be very efficient in stabilizing the effective charges with respect to geometry changes in all useful cases.

Another aspect of the use of eq. (3) is that it results in a systematic shift of the charges with respect to the use of eq. (1). This can be seen by considering the average q_{av} of the charges of a given atom over all conformers. The values of q_{av} obtained using the two equations are given in Table III. It is seen that they are similar for some atoms (difference smaller than 0.01 electron for H2, H1A1, H1A2, and H1A3), but that the difference can be larger than 0.2 electron for some others (C2A, C2, N3, and C3A). This shift is neither an advantage nor a disadvantage in regard to stabilization. It results from the way in which stabilization is obtained here and, furthermore, it makes the charges closer to the Mulliken charges and thus less dependent on the geometry (the variation in Mulliken charges with geometry is very small; see Table III). The shift also makes the charges closer to chemical intuition as discussed in the next section.

Qualitative Aspects

Some qualitative aspects of the charges obtained using eqs. (1) and (3) are worth mentioning. We consider here the charges obtained for conformer I using $scal = 2$.

We first look at atoms C1 and C2, which are in chemically similar situations; that is, they both belong to a carbonyl group and they are both binded to an aliphatic carbon on one side and to a NH group on the other side. Therefore, it can be reasonably expected that they have similar effective charges. Indeed, as seen in Table II (column *f*), the difference between the Mulliken charges of these two atoms is not very large (0.052 electron).

On the other hand, the difference between the charges obtained using eq. (1) is about three times larger (0.168 electron, see Table II, column *b*), and seems to indicate that the two atoms are chemically different for some reason. However, if we look at the charges obtained by solving eq. (3) we again find a small difference (0.028 electron), which means that fitting the electrostatic potential does

TABLE III.
Variation of Effective Atomic Charges in Alanine Dipeptide within the 12 Different Conformers Defined in Table I.^a

Atoms	q_{\min}	q_{\max}	$q_{\max} - q_{\min}$	q_{av}	q_{comp}
C2A	-.081	-.012	.069	-.051	
	-.020	.054	.074	.009	.002
	-.029	.621	.651	.281	.401
N2	-.884	-.816	.067	-.857	
	-.752	-.665	.086	-.712	-.713
	-.902	-.534	.368	-.739	-.845
H2	.370	.404	.034	.380	
	.306	.389	.083	.361	.350
	.285	.401	.115	.350	.358
C1	.724	.769	.045	.737	
	.883	.959	.075	.926	.917
	.760	.940	.180	.879	.951
O1	-.674	-.594	.080	-.622	
	-.761	-.674	.087	-.713	-.670
	-.732	-.658	.073	-.701	-.674
C2	.763	.821	.058	.799	
	.926	.967	.041	.944	.965
	.581	.824	.243	.684	.701
O2	-.644	-.606	.038	-.628	
	-.747	-.659	.088	-.703	-.692
	-.691	-.607	.084	-.657	-.650
N3	-.904	-.867	.036	-.876	
	-.746	-.623	.122	-.712	-.715
	-.502	-.367	.135	-.450	-.524
H3	.363	.450	.087	.383	
	.363	.495	.133	.407	.363
	.292	.382	.090	.339	.311
H2A	.206	.272	.066	.241	
	.099	.152	.053	.130	.111
	-.046	.163	.209	.051	-.013
C2B	-.471	-.396	.075	-.435	
	-.475	-.417	.057	-.449	-.450
	-.488	-.227	.260	-.366	-.227
H2B1	.134	.222	.088	.173	
	.129	.196	.067	.159	.154
	.041	.160	.119	.113	.064
H2B2	.141	.215	.074	.177	
	.096	.152	.056	.122	.129
	.132	.245	.113	.184	.047
H2B3	.132	.245	.113	.184	
	.077	.184	.107	.139	.153
	.046	.169	.123	.109	.073
C1A	-.537	-.529	.008	-.533	
	-.528	-.505	.023	-.516	-.524
	-.524	-.402	.122	-.474	-.603
H1A1	.214	.225	.012	.220	
	.151	.169	.018	.161	.147
	.138	.165	.028	.152	.165
H1A2	.175	.192	.018	.183	
	.123	.154	.031	.135	.144
	.106	.146	.040	.127	.166

TABLE III.
(continued)

Atoms	q_{\min}	q_{\max}	$q_{\max} - q_{\min}$	q_{av}	q_{comp}
H1A3	.172	.195	.023	.184	
	.112	.148	.036	.129	.141
	.092	.143	.052	.124	.165
C3A	-.203	-.195	.008	-.199	
	-.214	-.178	.036	-.198	-.196
	-.559	-.475	.083	-.511	-.366
H3A1	.178	.203	.025	.190	
	.115	.139	.024	.128	.121
	.189	.215	.025	.204	.161
H3A2	.152	.164	.013	.159	
	.096	.133	.037	.121	.138
	.176	.198	.022	.190	.172
H3A3	.182	.202	.020	.191	
	.114	.147	.033	.130	.125
	.193	.216	.023	.206	.165

^a For each atom, the first line refers to the Mulliken charges, the second line to the charges obtained using eq. (3), and the third line using eq. (1). q_{\min} = smallest charge of the different atoms among the 12 conformers; q_{\max} = largest charge; q_{av} = average of the charges in the different conformers; q_{comp} = composite charges fitting the potential of all the conformers.

not require consideration that the two atoms are in significantly different situations.

A similar conclusion is reached for atoms N2 and N3: they share a similar chemical environment, so similar charges could be expected. Indeed, the differences between the Mulliken charges (0.041 electron) and between the charges obtained by eq. (3) (0.062 electron) are rather small compared with the difference between the charges obtained using eq. (1) (0.184 electron).

The case of the carbon atoms in the two terminal methyl groups (C1A and C3A) is also worth discussing: the first atom, C1A, is bonded to a carbon atom (C1), and the second atom, C3A, to a nitrogen atom (N3). Therefore, it could be expected that C1A accommodates more electrons than C3A, because the carbon atom bonded to C1A attracts the electrons less efficiently than the nitrogen atom bonded to C3A. Indeed, this is the case resulting from the Mulliken charges: -0.533 electron for C1A and -0.199 electron for C3A.

The situation seems to be just the opposite when considering the charges resulting from eq. (1): -0.422 electron for C1A and -0.503 electron for C3A. Again, the chemical picture corresponding to the Mulliken charges is restored when using eq. (3): -0.517 electron for C1A and -0.193 electron for C3A.

The weak variation with geometry is another important qualitative aspect of the charges resulting from eq. (3) (see previous section).

On the whole, we have seen that the use of eq. (3) makes the charges generally (but not always) closer to the Mulliken charges than eq. (1). But, a perfect similarity is not necessarily desirable; the Mulliken charges usually have a reasonable chemical meaning, but they lead to a very poor fit of the potential (see Table II, column *f*, last two lines). Using eq. (3) leads to charges that are chemically meaningful, as just discussed, and still provide a good fit of the potential.

Composite Charges

We now assess whether it is possible to define charges that are independent of the geometry. To that end, we have fitted a unique set of “composite” charges on the values of the *ab initio* electrostatic potential at all fit points of all 12 conformers (i.e., at a total of 16,478 points—about 1400 points for each conformer). The points were determined using $scal = 2$. The fit was performed using either eq. (1) or eq. (3). The resulting charges are presented in Table III (column q_{comp}) and the

RMS and RRMS deviations in Table IV (columns *e* and *f*).

Qualitatively, the results are similar to those ones obtained by fitting charges separately for the different conformers:

- the fit obtained using eq. (3) is only slightly less accurate than that obtained by eq. (1); and
- the charges obtained by the two equations are very different, and the ones obtained by eq. (1) present several qualitatively unexpected aspects.

Thus, it is seen in Table IV that the composite charges lead to an average RMS deviation equal to 0.8 kcal/mol using eq. (3) and to 0.76 kcal/mol using eq. (1). These two RMS deviations are very close to each other, and they might be sufficiently low for some applications. It is also noteworthy that the RMS deviations evaluated separately for the different conformers using the composite charges are not much smaller than 1 kcal/mol, reaching 1.32 kcal/mol using eq. (3) for conformer III. The corresponding RRMS deviations are close to 8% (except for conformer III where it reaches 15%).

TABLE IV.
Root-Mean-Square Deviation (RMS) and Relative Root-Mean Square-Deviation (RRMS) between *Ab Initio* Electrostatic Potential and Potential Due to Effective Atomic Charges for 12 Conformers of Alanine Dipeptide.^a

Conformer	Computational method					
	Specific ^b		Average ^c		Composite ^d	
	Eq. (3)	Eq. (1)	Eq. (3)	Eq. (1)	Eq. (3)	Eq. (1)
	<i>a</i>	<i>b</i>	<i>c</i>	<i>d</i>	<i>e</i>	<i>f</i>
I	0.33	0.29	0.74	0.68	0.67	0.64
	4.4	3.8	9.7	8.9	8.9	8.4
II	0.34	0.29	0.80	0.78	0.66	0.64
	4.5	3.8	10.4	10.2	8.6	8.3
III	0.37	0.31	1.26	1.15	1.32	1.28
	4.3	3.6	14.5	13.2	15.2	14.7
IV	0.38	0.32	0.81	0.76	0.63	0.59
	4.0	3.3	8.4	7.9	6.5	6.2
V	0.32	0.25	0.98	1.03	0.82	0.79
	2.8	2.1	8.6	9.0	7.1	6.9
VI	0.40	0.33	1.98	1.901	0.83	0.77
	3.2	2.6	15.7	15.1	6.6	6.2
VII	0.30	0.26	1.08	1.10	0.72	0.70
	2.7	2.4	9.8	10.0	6.6	6.4
VIII	0.38	0.29	0.73	0.74	0.72	0.65
	4.3	3.3	8.2	8.3	8.1	7.3
IX	0.31	0.27	1.02	1.12	0.70	0.70
	2.4	2.0	7.9	8.6	5.4	5.4
X	0.40	0.32	2.05	1.98	0.86	0.81
	3.1	2.5	16.2	15.5	6.8	6.4
XI	0.33	0.30	0.64	0.61	0.69	0.68
	4.7	4.2	9.1	8.7	9.9	9.7
XII	0.32	0.28	0.71	0.68	0.72	0.71
	4.4	3.8	9.7	9.3	9.7	9.6
average			1.16	1.13	0.80	0.76
			11.6	11.3	8.0	7.6

^a For each conformer, the first line gives the RMS (kcal/mol) and the second line the RRMS (%). The last two lines ("average") give the RMS and the RRMS obtained by averaging the deviations over all 12 conformers. The column headings correspond to the methods used to evaluate the charges.

^b Charges obtained by fitting the electrostatic potential in each specific conformer either by means of eq. (1) (column *b*) or by means of eq. (3) (column *a*).

^c Average of specific charges of the 12 conformers.

^d Charges obtained by fitting the electrostatic potential of the 12 conformers simultaneously.

These errors are significantly larger than the ones obtained when fitting charges separately for the different conformers (around 0.3 kcal/mol, or 3–4%), but they are better than those obtained using average charges (about 1.1 kcal/mol or 11% with both equations).

Concerning the charges themselves, it is seen in Table III that the composite charges obtained by using eq. (1) present the same surprising aspects for C1 and C2, for N2 and N3, and for C1A and C3A, as discussed in the previous section. In addition, the composite charges obtained using eq. (1) are outside the range of charges specific for different conformers of several atoms (C1, N3, H2B2, C1A, H1A2, H1A3, C3A, H3A1, and H3A3). The composite charges obtained using eq. (3) are significantly different from those obtained using eq. (1), particularly for atom C2A. In addition, they have the same qualitative agreement with the Mulliken charges and with the "chemical intuition" (as discussed earlier) as the specific charges for atoms C1, C2, N2, N3, C1A, and C3B, and they lie within the range of specific charges (except for H3A2, where it is outside that range by 0.005 electron).

We conclude that the composite charges obtained by eq. (3) are more satisfactory than those obtained by eq. (1), in exactly the same way as the specific charges. However, the accuracy of the fit provided by the composite charges might be not sufficient in some cases. A mixed approach using specific charges for some atoms and composite charges for the remaining atoms could be then useful.

Conclusion

We have used a recent and original method for determining effective atomic charges in the case of alanine dipeptide. This method consists of fitting an *ab initio* electrostatic potential and in using a specific technique to ensure numerical stability of the fit [eq. (3)].

We have shown that the resulting charges fit the potential accurately, while retaining a reasonable chemical meaning (e.g., similar charges for atoms in a similar environment; see Qualitative Aspects section). In fact, these charges are close to Mulliken charges; they are necessarily slightly different from the Mulliken charges because the latter are unable to reproduce correctly the electrostatic potential. In addition, they do not require a com-

plete revision of the chemical idea of atomic charges, unlike the charges obtained by nonstabilized methods [like eq. (1)] and yet they ensure a good fit of the electrostatic potential, a characteristic not provided by Mulliken charges.

In regard to interatomic forces, it is likely that the two types of charges (obtained by using a stabilized or a nonstabilized method) are equivalent, because they both provide a good fit of the potential. The situation might be different in regard to determination of system conformation, because the two types of charges are significantly different and that the fit processes result in appropriate compensation only outside of the system. In such cases, it is to be expected that the charges obtained with a stabilized method are better because they are chemically much more reasonable (as discussed earlier). Nonetheless, this remains to be tested.

We have also tried to define a set of charges independent of the geometry of the system. The quality of the fit is therefore much less satisfactory (RMS deviation equal to nearly 1 kcal/mol). In the cases in which such accuracy is not sufficient, it is necessary to fit the geometry dependence of the charges. Such a fit is probably rather difficult with nonstabilized charges, because their geometry dependence appears to be rather erratic. A fit should be much easier with the stabilized charges, because their variations with geometry are very limited and smooth.

Acknowledgments

The authors are indebted to J. Smith (CEA, Saclay) and to P. Pernot (LPCR, Orsay) for useful discussions concerning this work.

References

1. Wiberg, K. B.; Rablen, P. R. *J Comput Chem* 1993, 14, 1504.
2. Bacrach, S. In *Reviews in Computational Chemistry*; Lipkowsky, K. B.; Boyds, D. B., Eds.; VCH: New York, 1994, p 171.
3. Sigfridsson, E.; Ryde, U. *J Comput Chem* 1998, 19, 377.
4. Stouch, T. R.; Williams, D. E. *J Comput Chem* 1992, 13, 622.
5. Williams, D. E. *Biopolymers* 1990, 29, 1367.
6. Spackman, M. A. *J Comput Chem* 1996, 17, 1.
7. Stouch, T. R.; Williams, D. E. *J Comput Chem* 1993, 14, 858.
8. Francl, M. M.; Carey, C.; Chirlian, L. E.; Gange, D. M. *J Comput Chem* 1996, 17, 367.

9. Dinur, U.; Hagler, A. T. *J Comput Chem* 1995, 16, 154.
10. Lévy, B.; Enescu, M. *J Mol Struct (Theochem)* 1998, 432, 235.
11. (a) Delhommelle, J.; Granucci, G.; Millié, P.; Boutin, A.; Fuchs, A. H. *J Comput Chem*, in press. (b) Lévy, B.; Rabbe, C.; Dognon, J. P.; Durand, S.; Granucci, G.; Millié, P.; Delhommelle, J., in press.
12. Schmidt, M. W.; Baldridge, K. K.; Boatz, J. A.; Elbert, S. T.; Gordon, M. S.; Jensen, J. H.; Koseki, S.; Matsunaga, N.; Nguyen, K. A.; Su, S. J.; Windus, T. L.; Dupuis, M.; Montgomery, J. A. *J Comput Chem* 1993, 14, 1347.
13. Price, S. L.; Andrews, J. S.; Murray, J. S.; Amos, R. D. *J Am Chem Soc* 1992, 114, 8268.